Accident Analysis and Prevention 73 (2014) 399–411

Contents lists available at ScienceDirect

Accident Analysis and Prevention

journal homepage: www.elsevier.com/locate/aap

Emergency healthcare worker sleep, fatigue, and alertness behavior survey (SFAB): Development and content validation of a survey tool

P. Daniel Patterson [a,*], Daniel J. Buysse [b], Matthew D. Weaver [a], Brian P. Suffoletto [a], Kyle L. McManigle [a], Clifton W. Callaway [a], Donald M. Yealy [a]

[a] Department of Emergency Medicine, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA
[b] Department of Psychiatry and Clinical Translational Science, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA

ARTICLE INFO

Article history:
Received 5 May 2014
Received in revised form 11 August 2014
Accepted 30 September 2014
Available online 15 October 2014

Keywords:
Sleep
Fatigue
Alertness
Emergency workers

ABSTRACT

Background: Workplace safety is a recognized concern in emergency medical services (EMS). Ambulance crashes are common and injury rates exceed that of the general working public. Fatigue and sleepiness during shift work pose a safety risk for patients and EMS workers. Changing EMS worker behaviors and improving alertness during shift work is hampered by a lack of instruments that reliably and accurately measure multidimensional beliefs and habits that predict alertness behavior.
Objectives: We sought to test the reliability and validity of a survey tool (the sleep, fatigue, and alertness behavior survey [SFAB]) designed to identify the cognitions of EMS workers concerning sleep, fatigue, and alertness behaviors during shift work.
Methods: We operationalized the integrative model of behavioral prediction (IMBP) and developed a pool of 97 candidate items and sub-items to measure eight domains of the IMBP. Five sleep scientists judged the content validity of each item and a convenience sample of EMS workers completed a paper-based version of the SFAB. We retained items judged content valid by five sleep scientists and performed exploratory factor analysis (EFA), confirmatory factor analysis (CFA), and tests of reliability and internal consistency. We identified a simple factor structure for each scale and calculated means and standard deviations for each item and scale.
Results: We received 360 completed SFAB surveys from a convenience sample of 800 EMS workers attending two regional continuing education conferences (45% participation rate). Forty-seven candidate items and sub-items/options were removed following content validation, EFA, and CFA testing. Analyses revealed a simple factor structure for seven of eight domains and a final pool of 50 items and sub-items/options. Domains include: attitudes, normative beliefs, knowledge, salience, habits, environmental constraints, and intent. EFA tests of self-efficacy items failed to identify a simple factor structure. We retained two self-efficacy items based on Spearman–Brown correlation of 0.23 ($p < 0.0001$).
Conclusions: Measurement of sleep, fatigue, and alertness behavior among EMS workers is challenging. We describe the development and psychometric testing of a survey tool that may be useful in a variety of applications addressing sleep, fatigue, and alertness behavior among EMS workers.
© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

1.1. Fatigue and EMS worker safety

Fatigue is an "unpleasant symptom incorporating feelings of tiredness to exhaustion creating conditions (physical and mental) that interfere with the ability to function in a normal capacity" (Ream and Richardson, 1996). Occupational fatigue has been linked to performance and negative safety outcomes in many occupations (Williamson et al., 2011). Emergency medical services (EMS) is a health care sector where a high proportion of workers self-report high-levels sleepiness and fatigue while on duty. Greater than 50% of EMS workers report high levels of mental and physical fatigue during shift work (Patterson et al., 2010a, 2012c). Compared to non-fatigued EMS workers, EMS workers reporting excessive fatigue also report poor sleep quality and problems with daytime sleepiness (Patterson et al., 2010a, 2012c; Pirrallo et al., 2012). Recent research shows that fatigue is a significant safety risk for

* Corresponding author at: Department of Emergency Medicine, University of Pittsburgh School of Medicine, 3600 Forbes Avenue, Iroquois Bldg., Suite 400A, Pittsburgh, PA 15261, USA. Tel.: +1 412 647 3183; fax: +1 412 647 6999.
E-mail address: pattersond@upmc.edu (P. D. Patterson).

http://dx.doi.org/10.1016/j.aap.2014.09.028
0001-4575/© 2014 Elsevier Ltd. All rights reserved.

EMS worker safety. Odds of injury are 1.9 times greater among fatigued EMS workers than the non-fatigued, after controlling other variables (Patterson et al., 2012c). Despite these data, and the obvious potential for worker and patient harm, we have a limited understanding of EMS worker fatigue and injury (Patterson et al., 2012b). A critical step to improving our understanding and intervening is identifying the behavioral factors that may be connected to fatigue and alertness in EMS. In this paper, we describe an initial development of reliable and valid measures of behavioral factors associated with EMS worker alertness.

EMS personnel work in shifts, operate ambulances at high-speeds, and deliver life-saving care prior to transporting critically ill and injured patients to the hospital (Brice et al., 2012). EMS work is often described as unpredictable, stressful, and dangerous. EMS workers face numerous threats to personal safety and the safety of their patients (Brice et al., 2012). Transportation of patients in ambulances is a key function of EMS operations and ambulance crashes are a recognized safety problem (Drucker et al., 2013; Ray and Kupas, 2007). The true magnitude of ambulance crashes and harm to patients and workers nationwide is unclear, yet believed to be more severe than that measured in other occupations and healthcare settings (Reichard et al., 2011). Worker fatigue is believed to be a key factor contributing to negative safety outcomes for patients and EMS workers (Patterson et al., 2012b).

Reducing EMS worker fatigue and sleepiness during shift work is a national priority (Bowman et al., 2013; NEMSAC, 2013; NHTSAOEMS, 2013). Employers, professional societies and organizations, insurers, the U.S. National Institute for Occupational Safety and Health (NIOSH), and the U.S. National Highway Traffic Safety Administration's (NHTSA) Office of EMS support improving safety by addressing EMS worker fatigue. Strategic goal 15.1 of the National Occupational Research Agenda (NORA) identifies worker fatigue as a problem for the EMS industry (Bowman et al., 2013). In 2013, the National EMS Advisory Council (NEMSAC) – a council that advises the federal government on matters related to EMS – recommended the Office of EMS within NHTSA to investigate gaps in knowledge, standardize investigations involving ambulance crashes, and disseminate information to local EMS employers (NEMSAC, 2013).

Fatigue among EMS workers is a visible issue at the national and federal level, yet efforts to address it are limited. Most often, employers of shift workers take it upon themselves to address the threat of fatigue by first reducing or eliminating extended shift lengths or overnight shifts (Caruso et al., 2004, 2006; Dawson et al., 2012). Shift work is commonly described as "any arrangement of work hours other than standard daylight hours" (IARC, 2010). Common shift work patterns in EMS include 12 and 24-h shifts (Patterson et al., 2010c, 2012a). There is no standardized shift pattern or length of shift in EMS, and no universal limit on how long an EMS worker can be on duty. Most EMS personnel work >45 h per week, and 40% work >16-shifts per month (Patterson et al., 2010c, 2012c). Many work multiple EMS jobs and amass unsafe amounts of overtime (Bauder, 2012; Patterson et al., 2010c, 2012c). One-third of EMS workers report excessive daytime sleepiness and deficits in performance due to their sleepiness, such as problems remembering protocols and difficulty operating an ambulance (Pirrallo et al., 2012). Some in the EMS industry advocate for the termination of extended shift work (e.g., 24-h shifts), referring to them as "unsafe" (Kirkwood, 2012). There is a substantial literature and body of research on shift work, with data showing that shift work disrupts the natural physiological regulation of sleep (Borbely, 1982; Daan et al., 1984). Research in non-EMS settings shows that extended shifts can lead to prolonged periods of wakefulness, mental and physical fatigue, increase an individual's drive for sleep, reduce alertness, and impair performance (Borbely, 1982; Daan et al., 1984).

Eliminating extended shifts or specific patterns of shift work may not remove or reduce the problem of fatigued EMS workers (Dawson et al., 2012). Many EMS workers are employed by multiple organizations – in some locations greater than 80% work for more than one EMS agency (Frakes and Kelly, 2007). From an EMS worker's perspective, working multiple jobs or accumulating overtime may be perceived as necessary in order to achieve a desired standard of living (Bauder, 2012; Greene, 2013). In 2012, half of all compensated EMS workers made less than $31,020 annually, or less than $14.91 per hour (Bureau Of Labor Statistics, 2014). Organizations in rural areas staff their ambulances with volunteers whom have full-time or part-time jobs outside of EMS. Odds are high that an EMS worker will show up to work fatigued or become fatigued during the course of a shift, which may or may not be due to extended shifts used by EMS employers.

To address fatigue in the workplace, sleep and fatigue specialists advocate that safety management systems address multiple factors simultaneously and integrate both fatigue-reduction and fatigue-proofing interventions (Dawson et al., 2012). Fatigue-reduction refers to policies that reduce the odds that a worker arrives to work fatigued. Examples include restricting shift length and specifying a maximum number of hours or days of shift work. Fatigue-proofing refers to strategies used while at work to reduce likelihood of a negative safety outcome due to fatigue. Examples include safety checklists, closed-loop communication, and use of technology that monitors performance and warns the worker of pending danger.

1.2. Current research and research gaps

It is unclear if fatigue-reducing and/or fatigue-proofing interventions or policies will impact EMS worker fatigue during shift work and subsequently reduce risk of negative safety outcomes. One study by Boudreaux et al. (1998) involved testing a fatigue-reduction intervention (cutting shift length) and showed lower perceived burnout after reducing shifts length from 24 to 12 h, yet this effect was short-term and did not have a long-term impact [12-months after implementation]. There is limited research involving EMS clinicians testing the impact of fatigue-proofing strategies on EMS worker perceived fatigue or patient outcomes. One study of 10 paramedics in Japan showed that paramedics provided scheduled naps had no impact on reaction time or self-reported fatigue compared to paramedics with no nap (Takeyama et al., 2009). Present-day use of fatigue-reducing or fatigue-proofing strategies in the EMS setting would be based on the generalization of evidence from non-EMS settings. This poses a problem for many in EMS that perceive the nature of work to be different from other shift work occupations where work is predictable and less stressful.

Sleep, fatigue, and shift work interactions are well studied in other professions (Belenky and Akerstedt, 2011; Caruso et al., 2004). We believe a line of research should determine why an EMS worker decides to work fatigued or considers it acceptable to work while severely fatigued and in a state of reduced alertness.

1.3. Theoretical framework and study objectives

Theoretical frameworks map complex relationships between factors associated with behavior and help explain why an individual chooses to engage in or fail to adopt a given behavior. The theory of reasoned action (TRA) and theory of planned behavior (TPB) are two leading theories of health behavior (Fishbein and Yzer, 2003). The integrative model of behavioral prediction (IMBP) is an amalgamation of the TRA and TPB and a principal theoretical model of behavior prediction (Fig. 1) (Fishbein and Yzer, 2003). Core determinants of behavior in the IMBP

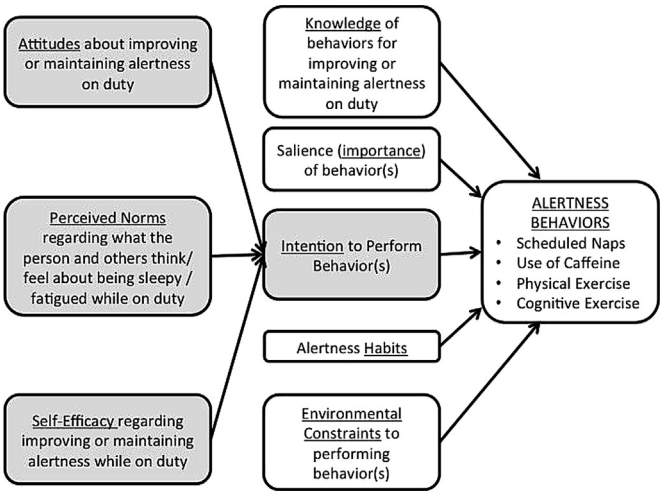

Fig. 1. Conceptual model.

include: (1) an individual's attitude concerning the behavior or behaviors and whether or not his/her opinions about behavior(s) are positive or negative; (2) an individual's perceived norms of what his/her peers may think of the behavior(s); and (3) an individual's self-efficacy or his/her belief that he/she can perform the behavior(s) in question. The IMBP posits that an individual will perform a behavior (e.g., coming to work alert and not fatigued) when he/she has: (a) formed a strong intention to perform the behavior, (b) has the essential skills to perform the behavior, and (c) constraints on the behavior are limited. Investigators can determine why an individual is or is not performing a specific behavior or behaviors by obtaining measures of beliefs, attitudes, perceived norms, self-efficacy, intention, habits, and environmental constraints (Fishbein and Yzer, 2003). Isolating the strongest predictors of a given behavior or behavioral intent improve odds of changing behavior with targeted and tailored interventions.

In our study, we sought to develop a preliminary set of reliable and valid scales to measure predictors of alertness behaviors among EMS shift workers. It is our intent to set the groundwork for additional investigations of alertness, fatigue, and sleep among EMS shift workers.

2. Methods

2.1. Overview of study design and study sample

We used a cross-sectional survey design including sleep scientists and EMS workers. We define EMS workers as those currently certified and actively working at the emergency medical technician (EMT) basic, intermediate, or paramedic level. Nurses certified as EMS air-medical flight nurses, and EMS physicians currently working as medical directors of EMS agencies were also considered EMS workers. We define 'sleep scientists' as doctoral-level clinical scientists with a record of academic publication in sleep, fatigue, or shift work.

We recruited EMS workers from western Pennsylvania, United States. In 2013, the regional EMS council confirmed a total EMS worker population of 13,300 currently registered with the state of Pennsylvania and certified to work in the western region of Pennsylvania. We recruited from this population using convenience sampling at two regional emergency medicine conferences hosted in western Pennsylvania in February and March 2014. Table 1 shows that the demographic characteristics of our study sample are analogous to samples of other studies using both convenience and random sampling of EMS workers.

Table 1
Demographic characteristics of EMS workers.

	Study sample EMS workers $N = 355$
Mean age (std.)	41 (12)
Sex	
Male	58%
Female	42%
Certification/license	
EMT-basic	43%
Paramedic	32%
Nurse	17%
Physician	1%
Other	7%
Clinical setting	
Ground-EMS	62%
Air-medical	3%
ED/ICU setting	20%
Other	15%
Years of experience	
Mean (std.)	16 (11)
Numb. of Jobs	
≥2	46%
Employment status	
Full-time	59%
Part-time	20%
Volunteer	21%
Number of shifts worked last month	
Mean (std.)	13 (10)
Most common shift length worked	
24-h	20%
12-h	31%
<12-h/other	49%
General health	
Excellent	22%
Good	70%
Fair/poor	8%
BMI	
Normal weight	18.7%
Overweight	29.8%
Obese	51.5%
Cigarettes/Wk	
Mean (std.)	78 (52)
Don't smoke	83%
Alcohol drinks/Wk	
Mean (std.)	4.3 (4.2)
Don't drink	43%
Diagnoses	
Diabetes	10%
Hypertension	27%
Heart problems	5%
Sleep apnea	11%
Breathing problems	8%
Arthritis	11%
Weight problems	31%
Migraines	12%
Depression	10%
Other	7%
Any condition	62%
Sleep quality	
Mean PSQI	6.7 (3.4)
% PSQI ≥6	27%
CFQ fatigue scale	
% Fatigued	64%
ESS Results	
Mean ESS (std.)	7.3 (5.3)
ESS ≥16	8%
ESS 10–15	23%
ESS 8–9	13%
ESS 0–7	56%
Mean KSS (std.)	4 (2.2)
Mean fatigue now (std.)	4.1 (2)
OFER results	
Chronic fatigue	40 (24.3)
Acute fatigue	51 (20)
Inter-shift fatigue	46 (18.9)
Mean SAS score	43 (22)

We then used convenience sampling to recruit a minimum of five sleep scientists. We recruited sleep scientists to follow established procedures and provide content validation ratings of our draft survey tool (Lynn, 1986). We selected sleep scientists based on a review of peer-reviewed publications on sleep related issues and recruited five scientists with >50 publications indexed in the PubMed/Medline databases. This study was approved by the University of Pittsburgh IRB.

2.2. Survey measures

EMS workers answered a 171-item paper survey administered in-person by the study's principal investigator. The survey was anonymous and participation was completely voluntary. The first 74 items included 14 demographic items; the widely used Pittsburgh Sleep Quality Index [PSQI, 18 items] (Buysse et al., 1989); the Epworth Sleepiness Scale [ESS, 8 items] (Johns, 1991); the Chalder Fatigue Questionnaire [CFQ, 11 items] (Chalder et al., 1993; Patterson et al., 2010a, 2012c); items from the general affect sub-scale and social/family sub-scale of the Schedule Attitudes Survey [SAS, 5 items] (Dunham and Pierce, 1986); the Occupational Fatigue, Exhaustion, Recovery Scale [OFER, 15 items] (Winwood et al., 2005); the Karolinska Sleepiness Scale [KSS, 1 item] (Akerstedt and Gillberg, 1990); and a new and previously untested "Fatigue Now" item analogous in form to the KSS and developed by investigators of this study [FN, 1 item].

We used the IMBP as a framework for developing pilot measures of factors key in the prediction of EMS worker alertness behavior and intent to adopt alertness promoting behaviors (Fishbein and Yzer, 2003). We developed 97 candidate items (4 items with multiple options or sub-items) intended to measure eight domains of the IMBP; these items are hereafter referred to as the sleep, fatigue, and alertness behavior survey (SFAB). Specifically, the SFAB measures the following domains:

- Attitudes towards maintaining alertness and reducing fatigue at work (ATT scale, 20 items). We define attitudes as a person's perceptions of alertness behavior(s) in terms of being positive/favorable or negative/unfavorable. We examined previous research by Ganley and Rosario (2013) that utilized 12 items to measure attitudes towards smoking among young adults. We adapted several items used by Ganley and Rasario to fit our purposes: (i.e., The item, "I do not like being around others who smoke" was edited to read: I do not like being around co-workers that try to stay alert for the entire shift.). Respondents recorded their responses with a six-point Likert scale with no neutral response option anchored with strongly disagree to strongly agree. All items were negatively worded and required recoding to ensure directional consistency where lower scores indicate negative attitudes.
- Normative beliefs (NB scale, 2 item stems, each with 5 sub-items/options, $n = 10$ total). We define normative beliefs as a person's perceived social pressure to perform or not perform behavior(s) of interest. Beliefs about what others think one should do, and beliefs about what a person's social networks are doing with respect to the behavior(s). Candidate items for this domain were developed using prior research by Olds et al. (2005) who investigated smoking and substance abuse among teenagers. We adapted wording of items used by Olds et al. to fit our study topic and population (i.e., the item stem: "What would the following people think of you if you started cigarettes" was edited to read, "What would the following people think of you if you were very sleepy and fighting the urge to sleep while at work?"). We included five reference groups (options) to which respondents responded: (1) people your age; (2) co-workers; (3) other healthcare workers; (4) my supervisors; and (5) other shift workers. Respondents recorded their response for each reference group using a six point Likert scale with no neutral response option. The scale was anchored from strongly disapprove to strongly approve. For these items, response options (strongly disagree) were considered to be positive reflection of the construct measured. All items required recoding to ensure directional consistency where lower scores indicate negative normative beliefs.
- Self-efficacy to improve or maintain alertness and reduce fatigue while at work (SE scale, 13 items). Self-efficacy refers to a person's belief in his/her ability, capability, and control to perform specific behavior(s) (Bandura, 2006). "Perceived self-efficacy is a judgment of capability" (Bandura, 2006). Candidate items were developed using the item development approach prescribed by Bandura (2006). This approach involves use of a single stem "Rate your degree of confidence by recording a number from 0 to 10" followed by a list of behaviors a respondent then rates from 0 to 10 with 0 anchored as "cannot do at all," 5 anchored as "moderately can do," and 10 "highly certain can do." Example items include: (Use a regularly scheduled nap on duty _____ (0–10); control feelings of sleepiness on duty by resting at work_____(0–10); drink an energy beverage when sleepy_____(0–10); etc.).
- Knowledge of fatigue reducing behaviors (KNOW scale, 3 standard items and 1 additional item with 6 sub-items/options, $n = 10$ total). Knowledge of behaviors refers to the awareness of something based on acquisition of information, understanding, or skills through experience and/or education. We developed candidate items based on prior research by Delucchi et al. (2009) whom developed a scale to measure knowledge, attitudes, and practices related to smoking cessation and treatment. An example item includes: "The hazards of fatigue and sleepiness on duty have been clearly demonstrated." Respondents recorded their response for each item using a six point Likert scale anchored from strongly disagree to strongly agree with no neutral response option.
- Importance (salience) of alertness and fatigue reduction behaviors while at work (IMPORT scale, 11 items). The salience of alertness behaviors refers to the individual's perceived importance or prominence of the behavior(s). Draft items for assessing salience were developed using prior research, which focused on measuring the salience of one's career (Allen and Ortlepp, 2002). Allen and Ortlepp collated items from previous research and developed two scales, one for measuring career salience and one for measuring work salience. We adapted items (i.e., "The most important things I do in life involve my career" for our purposes to read: "The most important thing I do during my shift work is to maintain alertness and reduce fatigue." Respondents recorded their response for each item using a six point Likert scale anchored from strongly disagree to strongly agree with no neutral response option. Six items were negatively worded and required recoding to ensure directional consistency where lower scores indicate negative perceptions of the salience construct.
- Environmental constraints associated with performing alertness-improving and fatigue-reducing behaviors at work (EC scale, 7 standard items and 1 additional item with 9 sub-items/options, $n = 16$ total). We define environmental constraints as factors external to the individual that can act as an inhibitor or facilitator of behavior(s) that can impact alertness and fatigue during shift work. We developed our candidate items for the EC scale by reviewing prior research (Delucchi et al., 2009). We edited the Delucchi scale to read: "There are various reasons that might limit your ability to reduce feelings of fatigue and sleepiness while on duty. Please rate the importance of each listed below when working at your primary EMS job (0 = not at all

important, 1 = somewhat important, 2 = important, 3 = very important)". Example barriers provided for rating include: ("Lack of time provided by management; Lack of training how to do it; Lack of education materials"). All items are negatively worded, but did not require reverse coding given the meaning of the 4 point Likert response scale. Seven additional items required a response using the six point Likert scale anchored from strongly disagree to strongly agree with no neutral response option. One of these items was negatively worded and required recoding to ensure directional consistency with other items measuring the construct.

- Habits that may aid or inhibit adoption of behaviors that improve alertness and reduce fatigue at work (HABIT scale, 7 items). We define habits as a person's tendencies or practices that he/she performs on a usual/repeated basis related to alertness and fatigue while at work. Candidate items for the HABIT scale are based on previous research by De Pelsmacker and Janssens (2007) whom developed a set of scales to reliably and validly measure normative and descriptive norms, personal identity, habit formation, and other factors associated with speeding behaviors. We adapted items in the habit construct of their scales to fit our context and study population: (i.e., "I frequently drive faster than allowed by speed limits" to read as "I frequently show up to work fatigued"). Respondents recorded their response for each item using a six point Likert scale anchored from strongly disagree to strongly agree with no neutral response option. Three items were negatively worded and recoded to ensure directional consistency.
- Intention to perform behaviors that improve alertness (INTENT scale, 10 items). We define intention as a signal or indication of a person's readiness or choice to perform a behavior or behaviors of interest. Outside of observing actual behavior, measurement of intention is used as a best available predictor of actual behavior (e.g., intent to quit one's job/turnover; intent to quit smoking/smoking cessation) (Griffeth et al., 2000; Rise et al., 2008). We developed 10 candidate items designed to measure EMS worker intention to perform alertness-improving and fatigue-reducing behaviors at work by adapting items used previously by Ganley and Rosario (2013) to measure smoking cessation. We adapted items to fit our context and study population (i.e., "I smoke cigarettes now but intend to quit within the year" to read as "In the next six months I plan to adopt a new strategy to improve my alertness at work"). Respondents recorded their response for each item using a six point Likert scale anchored from strongly disagree to strongly agree with no neutral response option. Five items were negatively worded and recoded to ensure directional consistency where lower scores indicate negative intentions.

Sleep scientists were instructed to review the content of all 97 candidate items and sub-items/options of the draft SFAB and provide a rating of each item's relevancy using a scale prescribed by Lynn (1986). Content validity refers to how well the items of a scale represent the full domain or range of content a scale is intended/should be measuring (Nevo, 1985). It is recommended that experts, rather than laypersons or members of a targeted population, perform the adjudication of a measure's content validity (Lynn, 1986; Nevo, 1985; Rubio, 2005). A minimum of five experts with evident knowledge in a particular field is commonly deployed for this purpose (Lynn, 1986). Experts rated the relevancy of each item using an electronic survey tool and a 4-point scale: (1 = not relevant; 2 = somewhat relevant, 3 = quite relevant; 4 = highly relevant) (Lynn, 1986).

2.3. Study protocol

We recruited expert sleep scientists to rate the content validity of SFAB candidate items. We used a review of the pertinent sleep, shift work, and fatigue literature to identify content experts and to solicit participation. Recruitment included an IRB approved study flyer. A total of 5 experts were recruited, which includes one co-author (DJB), and 100% agreed to participate. Experts recorded their rating of content validity for the draft SFAB survey tool using an electronic data collection tool. We offered a coffee mug and an engraved coin with popular EMS symbols and the name of the principal investigator's research laboratory for each expert as the sole remuneration.

EMS workers were recruited to answer all 171 survey items in reference to their work in the EMS setting. We recruited EMS workers by positioning two investigators at an exhibit booth at two regional conferences in western Pennsylvania occurring in February and March 2014. Investigators greeted individuals approaching the booth, screened for eligibility, and distributed the anonymous paper survey. Eligibility criteria included: (1) active EMS clinician in the western region of Pennsylvania, (2) certified or licensed as an emergency medical technician (EMT) at the basic, paramedic, or flight paramedic level, a prehospital/emergency nurse, or an EMS physician that provides both online and off-line medical oversight of EMS clinicians; and (3) 18 years of age or older. Subjects received one coffee mug and a coin engraved with popular EMS symbols once we received a completed survey.

2.4. Data analysis

We calculated frequencies, percentages, means, and standard deviations to describe the demographic characteristics of our study sample (Table 1). For established tools (i.e., PSQI, ESS, CFQ, SAS, OFER, and KSS), we calculated summary scores as prescribed in the literature. For all the items of the draft SFAB survey tool, we weighted response options on a 0–100 scale. Negatively worded items were reverse coded.

We collated sleep scientists content validity scores and calculated the item content validity index (I-CVI) using techniques prescribed by Lynn (Lynn, 1986). The I-CVI is the proportion of experts that assign a rating of 3 = "quite relevant" or 4 = "highly relevant" to a single item. An I-CVI score $\geq$0.78 is evidence the item is content valid (Lynn, 1986).

We then excluded non-content valid items and performed a series of exploratory factor analyses (EFAs) within each scale to further reduce the total number of items per scale and identify cohesion between remaining items. We retained items with a $\geq$ 0.40 factor loading and used the Proportion of Variance test to determine the total number of factors to retain. The Proportion of Variance test is the eigenvalue for the factor of interest divided by the total eigenvalues of the correlation matrix. We chose the Proportion test over the eigenvalue one criterion and screen test per the criticisms and suggestions posed by Hatcher (1994). In some cases, values for the Proportion of Variance test can exceed 100% because some factors have negative eigenvalues. We examined construct validity with confirmatory factor analysis (CFA). We evaluated six established measures of model fit (construct validity) and two measures of internal consistency (reliability). These include the goodness of fit index (GFI), Bentler's comparative fit index (CFI), root mean square error of approximation (RMSEA), Bentler and Bonnett's non-normed fit index (NNFI), Bentler and Bonnett's normed fit index (NFI), standardized root mean square residual (SRMR), Cronbach's alpha coefficients, item-factor Pearson correlation coefficients. Established benchmarks for construct validity measures indicating good model fit include: SRMR $<$0.08, RMSEA $<$0.06, and GFI, CFI, NNFI, and NFI $\geq$0.9 (Anderson and Gerbing, 1988; Hatcher, 1994; Lance et al., 2006; Marsh et al., 1988). Established benchmarks for reliability measures include: Cronbach's alpha coefficient scores $\geq$0.70, item-scale Pearson correlations >0.40 (DeVellis, 1991; Ware and Gandek, 1998). We created scale-level scores by

summing the weighted scores of items within a given scale and dividing that sum by the total number of items. We anticipated producing multiple scale-level scores within a given construct following EFA and CFA analyses (e.g., multiple scale-level scores within the attitudes scale [ATT scale]). See Table 5 for recommendations for item scoring and generation of summary statistics for each scale/construct of the pilot SFAB survey.

3. Results

We received content validity ratings from five sleep scientists and completed surveys from 360 EMS workers. A total of 105 surveys were collected from conference location 1, and 255 from conference location 2. Conference organizers reported a total of 282 individuals registered for conference location 1, and 518 registered for conference location 2. The exact number of eligible/ineligible participants at each conference is not known. Because it is not possible to stratify conference attendees by EMS worker or non-EMS worker status, we conservatively estimate a rate of EMS worker participation of (360/800 = 45%). Five surveys were excluded with greater than 50% of items skipped or missing, leaving 355 EMS worker surveys for psychometric analysis.

The demographic characteristics of our EMS worker sample are reported in Table 1. Most survey participants are male, middle-aged, certified as an EMT-basic or paramedic, and work at a ground-based EMS operation. The characteristics of our EMS worker sample are similar to those reported in other studies (Table 3). Scores on standardized sleep and fatigue survey tools (i.e., the PSQI) are reported in Table 1 and are representative of scores reported in previous studies of EMS workers.

Five sleep scientists reduced the list of candidate items by rating the content validity of all candidate items and sub-items (Table 3). Twenty-nine candidate items and sub-items/options failed to meet the 0.78 benchmark for content validity. Sixty-eight candidate items and sub-items/options were considered content valid and retained for psychometric analysis, including EFA, CFA, and reliability testing.

Ratings of content validity by sleep scientists revealed 13 of 20 candidate items (65%) of the ATT scale as content valid (Table 4). EFA and CFA analyses of the ATT scale suggested a two-factor model with six-items in factor one (ATT one) and three-items in factor two (ATT two). The proportion of variance test for ATT one was 62% and 33% for ATT two. Construct validity scores for were acceptable: SRMR = 0.03, GFI = 0.95, CFI = 0.94, NNFI = 0.90, NFI = 0.93. The RMSEA score 0.09 exceeded the 0.06 cut point. Tests of reliability show acceptable internal consistency for both ATT one ($a = 0.82$) and ATT two ($a = 0.72$), and acceptable item-scale correlations (>0.40 cut-point). The mean composite/scale-level scores for the sub-scales were 60.1 (std. 20.3) for ATT one and 35.5 (std. 21.8) for ATT two. Mean scores for individual items of the final ATT scale are presented in Table 4.

Ratings of content validity by sleep scientists revealed both core candidate items (100%) of the NB scale, each with five sub-item/options, as content valid (Table 4). Psychometric analyses of the NB scale suggested two separate factors (NB one and NB two), one for each stem with five sub-items. The proportion of variance test for NB one and NB two exceeded 100%. Construct validity scores for NB one were acceptable: SRMR = 0.05, GFI = 0.92, CFI = 0.94, NNFI = 0.89, NFI = 0.94. The RMSEA score 0.19 exceeded the 0.06 cut point. Construct validity scores for NB two were acceptable: SRMR = 0.02, GFI = 0.98, CFI = 0.99, NNFI = 0.98, NFI = 0.99. The RMSEA score 0.08 exceeded the 0.06 cut point. Tests of reliability show acceptable internal consistency for both NB scale one ($a = 0.89$) and NB two ($a = 0.93$), and acceptable item-scale correlations (>0.40 cut-point). The mean composite/scale-level scores for the sub-scales were 59.3 (std. 21.1) for NB one and 63.5 (std. 20.9) for NB two. Mean scores for item of the final NB scale are presented in Table 4.

Content validity ratings identified four-items in the self efficacy scale (SE scale) for EFA and CFA analyses. Factor loadings for two of four content valid items failed to meet the 0.40 loading benchmark. EFA analyses therefore did not identify a simple factor structure. Standard tests of construct validity and reliability were not possible (i.e., proportion of variance and Cronbach's alpha tests). We were left with a possible two-item measure of self-efficacy (Table 4). The Spearman–Brown correlation coefficient for the two-item SE scale was 0.23 ($p < 0.0001$), a reliability measure considered appropriate for of two-item measures (Eisinga et al., 2013). Mean scores for the two retained items are reported in Table 4.

All four of the KNOW scale items, including one item with six-sub items/options ($n = 10$ total items), exceeded the 0.78 cut-point for I-CVI and qualify as content valid (Table 4). The proportion of variance test for KNOW one was 87% and 16% for KNOW two. Construct validity scores for two factor structure were acceptable: SRMR = 0.01, GFI = 0.96, CFI = 0.99, NNFI = 0.97, NFI = 0.98. The RMSEA score 0.08 exceeded the 0.06 cut point. Tests of reliability show acceptable internal consistency for both KNOW one ($a = 0.96$) and KNOW two ($a = 0.87$), and acceptable item-scale correlations (>0.40 cut-point). The mean composite/scale-level scores for the sub-scales were 75.6 (std. 21.1) for KNOW one and 78.0 (std. 18.3) for KNOW two. Mean scores for individual items of the final KNOW scale are presented in Table 4.

Five of 11 candidate items (45%) of the IMPORT scale were judged content valid. Psychometric tests revealed a one-factor solution with three-items (Table 4). The proportion of variance test for the IMPORT scale was >100%. Construct validity scores for single factor structure were acceptable: SRMR = 0.00, RMSEA = 0.00, GFI = 1.00, CFI = 1.00, NFI = 1.00. Tests did not produce an NNFI score. The Cronbach's alpha coefficient score was 0.62, slightly less than the desired $\geq$0.70 cut-point. Additional tests of reliability using item-scale Pearson correlations reveal acceptable reliability with all item-scale correlations exceeding the $\geq$0.40 cut-point (Ware and Gandek, 1998). The mean composite/scale-level score for the IMPORT scale was 63.8 (std. 17.0). See Table 4 for mean scores of individual items.

All but one sub-item of the 16 total items and sub-items/options of the EC scale were judged content valid by sleep scientists (Table 4). Psychometric EFA and CFA tests reveal three factors EC one, EC two, and EC three. The proportion of variance test for EC one, EC two, and EC three were 96%, 73%, and 38%, respectively. Construct validity scores for EC one were acceptable: SRMR = 0.02, GFI = 0.98, CFI = 0.98, NNFI = 0.97, NFI = 0.97. The RMSEA score 0.08 exceeded the 0.06 cut point. Construct validity tests were not performed for EC two given only two factors remained after construct and factor analysis testing. Construct validity scores for EC three were acceptable: SRMR = 0.04, GFI = 0.96, CFI = 0.95, NNFI = 0.81, NFI = 0.95. The RMSEA score 0.17 and NNFI of 0.81 did not meet specified cut points. Tests of reliability were positive and confirm internal consistency and high reliability. The Cronbach's alpha coefficient score for EC one was 0.86 and 0.82 for EC three. The Spearman–Brown test for reliability of EC two was 0.69 ($p < 0.0001$). Item-scale Pearson correlations reported in Table 4 reveal acceptable reliability of EC one and EC three exceed the $\geq$0.40 cut-point (Ware and Gandek, 1998). Mean composite/scale-level scores for the three sub-scales include EC one 63.4 (std. 21.8), EC two 46.9 (std. 27.2), and EC three 53.4 (std. 22.6). See Table 4 for mean scores of individual items.

Content validity ratings identified six of seven candidate items of the HABIT scale as content valid (Table 4). Psychometric tests revealed a one-factor model with three total items. The proportion of variance test for HABIT scale was >100%. Construct validity

Table 2
Comparison of study sample with previous research.

	Study sample	Previous research					
	EMS workers $N = 355$	EMT-teamwork survey study	LEADS sample	High response EMS sample	NHTSA workforce report	ROC study EMS agencies	Nat'l survey of EMS safety culture
Mean age in years	41	36.9	–	–	35	–	–
Sex							
Male	58%	72.3%	72.9%	71.8%	71–77%	–	73.2%
Female	42%	27.7%	27.1%	28.2%	23–29%	–	26.8%
Certification/ license							
EMT-Basic	43%	28.7%	58.1%	50.7%	72%	58.2%	19.4%
Paramedic	32%	61.4%	34.6%	49.3%	22%	34.2%	62.1%
Nurse	17%	–	–	–	–	–	8.0%
Physician	1%	–	–	–	–	–	–
Other	7%	9.9%	–	–	–	–	10.5%
Employment status							
Full-time	59%	71.7%	–	–	89%	–	77.6%
Part-time	20%	30.2%	–	–	11%	–	20.6%
Volunteer	21%	6.6%	–	–	–	–	1.8%

The EMT-teamwork survey study (Patterson et al., 2012c). The LEADS Sample (Brown et al., 2002). High response EMS sample (Patterson et al., 2010b). The NHTSA workforce report (Chapman et al., 2008). The ROC study EMS agencies (Davis et al., 2007). The National survey of EMS safety culture (Patterson et al., 2010a).

scores for one factor structure were acceptable: SRMR = 0.00, GFI = 1.00, CFI = 1.00, NFI = 1.00. Tests did not produce an NNFI score. Tests of internal consistency were positive, indicating a high-level of internal consistency with three items (Cronbach's coefficient alpha = 0.90). Item sub-scale Pearson correlations confirm acceptable reliability (Table 4). The mean composite/scale-level score for the HABIT scale was 45.9 (std. 24.4). See Table 4 for mean scores of individual items.

Five of 10 candidate items for the INTENT scale were judged content valid by sleep scientists. Psychometric EFA and CFA tests revealed a one-factor model with three total items (Table 4). The proportion of variance test for INTENT scale was 91%. Construct validity scores for single factor structure were acceptable: SRMR = 0.00, GFI = 1.00, CFI = 1.00, NFI = 1.00. Tests did not produce an NNFI score. Test of reliability confirm the three items possess acceptable internal consistency (Cronbach's coefficient alpha = 0.71). Item sub-scale Pearson correlations reported in Table 4 exceed the cut-point for indicating acceptable reliability. The mean INTENT scale composite/scale-level score was 65.3 (std. 18.2). Mean item-level scores can be found in Table 4.

4. Discussion

Alertness promoting and fatigue reduction interventions are needed to improve the health and safety of EMS workers and the safety of their patients. We developed a preliminary set of scales for the specific purpose of classifying EMS workers into categories of behavior and intent related to alertness while on the job.

Few studies have developed reliable and/or valid tools for understanding the predictors of behavior or behavioral intent related to sleep, fatigue, and alertness. We identified two studies developing sleep health scales for undergraduate college students (Knowlden et al., 2012; Robbins and Niederdeppe, 2014). We did not identify similar research involving shift workers or specifically EMS workers. Knowlden et al. (2012) developed draft items for predicting sleep behaviors of undergraduate students by operationalizing the theory of planned behavior (TPB). Investigators began with a qualitative study using feedback from 11 students. Six experts provided "face validity" of a draft survey tool before the final version of a 29-item instrument was established. A sample size of 197 undergraduate students provided data for psychometric tests. Authors report acceptable findings for reliability and construct validity with a total of four factors linked to sleep behavior ((1) attitude toward the behavior; (2) subjective norm; (3) perceived behavioral control; and (4) behavioral intent). In a separate study, Robbins and Niederdeppe (2014) sought to operationalize the IMBP theoretical model to distinguish between college students that intend to engage in healthy sleep behavior and those that do not. Investigators began with eliciting open-ended responses from a convenience sample of 31 college

Table 3
Progression of SFAB item and scale development.

Scale	Initial pool of items	Met criteria for content validity	Final factor construction after EFA and CFA analyses
Attitudes scale (ATT Scale)	N = 20 items with some items having multiple response options	$N = 13$ items	ATT one ($N = 6$ items) ATT two ($N = 3$ items)
Normative Beliefs Scale (NB scale)	$N = 2$ core items (each item has $n = 5$ options) [$N = 10$ items total]	$N = 2$ core items (each item has $n = 5$ options) [$N = 10$ items total]	NB one ($N = 1$ item with 5 options) NB two ($N = 1$ item with 5 options)
Self-Efficacy Scale (SE scale)	$N = 13$ items	$N = 4$ items	SE scale ($N = 2$ items)
Knowledge Scale (KNOW scale)	$N = 4$ items ($N = 1$ of these items has 6 options) [$N = 10$ items total]	N = 4 items (N = 1 of these items has 6 options) [N = 10 items total]	KNOW one ($N = 1$ item with 6 options) KNOW two ($N = 3$ items)
Importance/Salience Scale (IMPORT scale)	$N = 11$ items	N = 5 items	IMPORT ($N = 3$ items)
Environmental Constraints Scale (EC scale)	$N = 8$ items ($N = 1$ of these items with 9 options) [$N = 16$ items total]	N = 8 items (N = 1 of these items with 8 options) [N = 15 items total]	EC one ($N = 1$ item with 5 options) EC two ($N = 2$ items) *EC* three ($n = 4$ items)
Habits Scale (HABIT scale)	$N = 7$ items	$N = 6$ items	HABIT scale ($N = 3$ items)
Intentions Scale (INTENT scale)	$N = 10$ items	$N = 5$ items	INTENT scale ($N = 3$ items)
Total Items	$N = 97$	$N = 68$	$N = 50$

Table 4
Final item pool and scale structure of the SFAB.

Scale/sub-scale/items	Content validity (I-CVI scores)	Item-scale Pearson Corr	Mean-item score (std)
ATT-scale – ATT one[a]			
[QATT3] Being sleepy at work does not bother me [R]	0.80	0.73	63.2 (26.0)
[QATT4] Feeling physically or mentally fatigued at work does not bother me [R]	0.80	0.80	66.4 (26.2)
[QATT5] I'm not really bothered by feeling sleepy or fatigued while at work [R]	0.80	0.82	60.6 (28.4)
[QATT6] Being sleepy or fatigued at work is less of problem than other issues [R]	0.80	0.75	52.1 (27.7)
[QATT7] Being sleepy or fatigued at work does not impact my performance [R]	0.80	0.74	57.7 (29.3)
[QATT11] Employers should not have the ability to keep sleepy or fatigued workers from working a shift [R]	1.00	0.52	60.2 (30.7)
ATT-scale – ATT two[a]			
[QATT12] I intend to show up to work if I feel sleepy or fatigued [R]	0.80	0.73	34.4 (26.8)
[QATT13] If my employer needs me to work overtime or cover a shift, I will work even if I'm sleepy or fatigued [R]	1.00	0.82	36.7 (27.1)
[QATT14] If I need extra money, I will work overtime or an extra shift when I'm sleepy or fatigued [R]	1.00	0.84	35.7 (28)
NB-scale – NB one[b]			
What would the following people think of you if you were sleepy and fighting the urge to sleep while at work?			
[QNB22A] People your age [R]	0.80	0.60	59.2 (24.9)
[QNB22B] Co-workers [R]		0.67	55.2 (25.5)
[QNB22C] Other healthcare workers [R]		0.66	58.4 (24.2)
[QNB22D] My supervisors[R]		0.59	67.4 (26.2)
[QNB22E] Other shift workers [R]		0/64	56.1 (25.2)
NB-scale – NB two[b]			
What would the following people think if you were very fatigued mentally or physically while at work?			
[QNB23A] People your age [R]	0.80	0.86	63.0 (24.1)
[QNB23B] Co-workers[R]		0.92	60.2 (24.1)
[QNB23C] Other healthcare workers [R]		0.92	62.9 (23.0)
[QNB23D] My supervisors [R]		0.78	70.5 (24.0)
[QNB23E] Other shift workers [R]		0.90	61.0 (24.2)
SE-scale			
Rate degree of confidence from 0–10—0 = Cannot do at all–to–10 = highly certain can do		Spearman–Brown Correlation 0.23	
[QSEF28] Drink a coffee when sleepy or fatigued (__)	0.80		64.4 (37.0)
QSEF33] Walk, jog, or run on duty to fight sleepiness or fatigue (__)	0.80		45.5 (36.0)
KNOW-scale – KNOW one[a]			
Fatigue and sleepiness at work increase my risk of			
[QKNW47A] Making a medical error	1.00	0.92	75.3 (22.9)
[QKNW47B] Being injured		0.93	76.0 (22.0)
[QKNW47C] Being involved in an ambulance crash		0.84	75.3 (24.6)
[QKNW47D] Making a medication error		0.93	75.6 (23.4)
[QKNW47E] Making a mistake		0.95	76.1 (22.6)
[QKNW47F] Losing the ability to concentrate		0.91	75.4 (23.3)
KNOW-scale – KNOW two[a]			
[QKNW48] The hazards of fatigue and sleepiness on duty have been clearly demonstrated	0.80	0.90	74.5 (23.1)
[QKNW49] Lack of sleep at home increase the risk of a fatigue-related error or injury while at work	1.00	0.87	80.5 (18.1)
[QKNW50] The benefits of getting adequate sleep before shift work have been clearly demonstrated	0.80	0.92	78.7 (20.1)
IMPORT-scale[a]			
[QSIMP51] The most important thing I do during my shift work is to maintain alertness and reduce fatigue	0.80	0.80	62.9 (24.4)
[QSIMP56] I am willing to make sacrifices while at work to reduce feelings of fatigue and maintain alertness	0.80	0.70	63.3 (22.0)
[QSIMP61] I would do whatever is necessary to reduce feelings of fatigue while at work and improve my alertness on duty	0.80	0.76	65.3 (17.0)
EC-scale – EC one			
There are various reasons that might limit your ability to reduce feelings of fatigue and sleepiness while on duty. Please rate the importance of each listed when working at your primary EMS job 4-point Likert scale (0 = not at all important – to – 3 = very important)			
[QEVC65] Lack of interest from management	1.00	0.84	63.6 (27.4)
[QEVC66] Lack of company policies that address this issue	1.00	0.81	64.5 (26.3)
[QEVC67] My employer is not interested	0.80	0.81	62.1 (29.0)
[QEVC69] Other issues are more important to the company	1.00	0.74	63.4 (26.7)
[QEVC70] My co-workers and I are not given time to do things that reduce fatigue or sleepiness	0.80	0.80	63.2 (27.0)
EC-scale – EC two[a]			
[QEVC71] Reducing fatigue or sleepy workers is a priority of my supervisors/company administrators	1.00	Spearman–Brown Correlation 0.69	47.2 (28.7)
[QEVC72] My company/employer has policies or procedures that address fatigue or sleepiness	1.00		46.7 (30.5)
EC-scale – EC three[a]			
[QEVC74] My responsibilities at home prevent me from doing things to reduce my fatigue or sleepiness when I'm at work	1.00	0.81	54.4 (26.1)

Table 4 (*Continued*)

Scale/sub-scale/items	Content validity (I-CVI scores)	Item-scale Pearson Corr	Mean-item score (std)
[QEVC75] My obligations to my family inhibit me from getting the sleep I need to recover between my shifts	1.00	0.84	54.3 (25.9)
[QEVC76] I have multiple jobs that prevent me from getting the rest I need between shifts	1.00	0.73	48.1 (32.2)
[QEVC77] I have too many competing commitments and do not have time to get adequate sleep when at home and reduce feelings or sleepiness while at work	1.00	0.86	56.9 (27.8)
HABIT-scale[a]			
[QHAB78] I frequently show up to work feeling fatigued [R]	1.00	0.87	50.6 (27.1)
[QHAB79] Working while fatigued is something I do without thinking aboutit [R]	0.80	0.93	43.4 (26.2)
[QHAB80] I work while feeling fatigued because I do this all the time [R]	0.80	0.94	43.9 (27.0)
INTENT-scale[a]			
[QINT39] I am trying to improve my alertness at work and avoid feelingsleepy or fatigued	0.80	0.60	67.7 (21.1)
[QINT44] I do not plan to do anything to reduce my feeling sleepy onduty [R]	0.80	0.88	64.5 (23.7)
[QINT45] I do not plan to do anything while on duty to improve or maintainalertness [R]	0.80	0.89	63.5 (23.9)

Per recommendations by Hatcher (1994), items with incomplete data for any item in a scale were deleted from EFA and CFA analyses. Therefore, testing for the ATT-scale was performed on $n = 337$ completed surveys, $n = 347$ surveys for the NB-scale and SE-scale, $n = 342$ surveys for the KNOW-scale and IMPORT-scale, $n = 327$ surveys for the EC-scale, $n = 342$ surveys for the HABIT-scale, and $n = 343$ surveys for the INTENT-scale. [R] refers to the item responses being reverse coded.

[a] Response option anchored with strongly disagree to strongly agree.

[b] Response option anchored with strongly disapprove to strongly approve.

students and developed a draft survey tool to measure components of the IMBP. A draft tool with an unknown number of items was administered to a convenience sample of 361 college students on one college campus in the northeast U.S. Details on psychometric testing were not provided. The authors report that their measures of attitude and perceived control predicted college student's behavioral intent of healthy sleep.

Our tests of content validity, a critical early step in improving construct validity and developing high-quality measurement (Polit et al., 2007), are a strength of this study. Common standards of self-report scale development include: (a) a clear definition of each construct measured; (b) reference to the literature surrounding a construct of interest; (c) a description of how items were constructed; (d) content validation of item pools; (e) tests of reliability; (f) tests of construct validity; (g) a description of the target population for the instrument; (h) description of the sampled respondents representativeness; (i) description of construct scoring; and (j) summary statistics of measures of central tendency and dispersion for each construct (Norbeck, 1985). We believe the thoroughness of our study; meeting all of the above criteria including content validation, provides a high-level of confidence in our final item pool.

Another strength of our study, it is grounded in the dominant theoretical model of behavior change/prediction, the IMBP (Fishbein and Yzer, 2003). A vulnerability of measurement tools that are not driven by theory is they may be ineffective in detecting change in factors significant to the behavior or outcome of interest. We used the IMBP to operationalize a pilot set of items to measure the immediate antecedent of alertness behaviors for EMS shift workers. With further study, we have the capability to expose specific beliefs/factors that (1) may be modifiable and help change an individual's intention to adopt positive behaviors, and (2) help interventionists design programs to promote maintenance of positive behaviors and beliefs.

We anticipate that interventionists may use the SFAB to classify workers into distinct categories of behavior and intent regarding alertness while at work. Differentiating workers into categories will be useful because different interventions will be necessary based on unique beliefs, attitudes, perceived norms, self-efficacy, intention, habits, and environmental constraints. For example, some EMS workers may have no intention to engage in fatigue-reduction behaviors during shift work. Interventions for these workers should be directed at changing their attitudes, perceived norms, or self-efficacy. Other EMS workers may have formed the intention to reduce fatigue by adopting fatigue-reduction behaviors, yet they have not acted on their intent. These workers may benefit from an intervention directed at skill building or removal of environmental constraints. The measurement scales from this tool may also identify workers that have a strong intention to reduce their fatigue while at work and perform a particular behavior (e.g., take scheduled naps). For these workers, it may be beneficial to not intervene with a reduction strategy and more simply assist these workers in maintaining their habits and beliefs.

For investigators, the SFAB may aid in comparing the impact of experimental sleep and fatigue interventions. Investigations that involve infrequent events or long-term follow-up are often cost-prohibitive. Outcomes such as injury or error are difficult to measure and often too infrequent for use in studies using shorter time intervals. The SFAB offers an opportunity to detect impact in studies of shorter duration. Specifically, a worker's attitude, knowledge, normative beliefs, self-efficacy, habits, or perceived importance of a given behavior may change quickly. Use of the SFAB scale pre and post intervention would permit detection of change in one or more factors associated with behavioral change. See Table 5 for a copy of the final list of SFAB survey items and constructs with corresponding Likert response scales and instructions for scoring and interpretation.

Improving safety of emergency workers, especially EMS workers, is a national priority and endorsed by the recently completed National EMS Culture of Safety Strategy project funded by the National Highway Traffic Safety Administration and led by the American College of Emergency Physicians (www.emscultureofsafety.org). Our study represents formative work in operationalizing established theories of behavior germane to sleep, fatigue, and alertness behaviors among emergency workers. Our SFAB survey should undergo repeated evaluation and testing (Rubio et al., 2003). We envision these scales will be preliminary in nature given the nascent research on sleep and fatigue in EMS, and the likelihood our current understanding of the issues will evolve with increased research. These scales and items are preliminary, yet integral to future research devoted to measuring factors influential in emergency worker alertness behavior.

5. Limitations

Our study is limited by the convenience sampling. There are several limitations associated with convenience sampling that may

Table 5
Final list of the pilot SFAB survey tool with instructions for scoring and interpretation.

Scale/sub-scale/items	Likert response scale and scoring
ATT-scale – ATT one	[R] = reverse scoring
[QSFAB-1] Being sleepy at work does not bother me [R]	Strongly disagree = 100
[QSFAB-2] Feeling physically or mentally fatigued at work does not bother me [R]	Disagree = 80
[QSFAB-3] I'm not really bothered by feeling sleepy or fatigued while at work [R]	Slightly disagree = 60
[QSFAB-4] Being sleepy or fatigued at work is less of a problem than other issues [R]	Slightly agree = 40
[QSFAB-5] Being sleepy or fatigued at work does not impact my performance [R]	Agree = 20
[QSFAB-6] Employers should not have the ability to keep sleepy or fatigued workers from working a shift [R]	Strongly agree = 0
ATT-scale – ATT two	[R] = reverse scoring
[QSFAB-7] I intend to show up to work if I feel sleepy or fatigued [R]	Strongly disagree = 100
[QSFAB-8] If my employer needs me to work overtime or cover a shift, I will work even if I'm sleepy or fatigued [R]	Disagree = 80
[QSFAB-9] If I need extra money, I will work overtime or an extra shift when I'm sleepy or fatigued [R]	Slightly disagree = 60
	Slightly agree = 40
	Agree = 20
	Strongly agree = 0
ATT-scale scoring: summary scores for ATT one and ATT two is the sum of item scores on divided by total number of items. The composite ATT scale score is the sum of ATT one and ATT two divided by two	
ATT-scale score interpretation: higher scores indicate a more positive/favorable attitude towards maintaining alertness and reducing fatigue while at work	
NB-scale – NB one	
What would the following people think of you if you were sleepy and fighting the urge to sleep while at work?	[R] = reverse scoring
[QSFAB-10] People your age [R]	Strongly disapprove = 100
[QSFAB-11] Co-workers [R]	Disapprove = 80
[QSFAB-12] Other healthcare workers [R]	Slightly disapprove = 60
[QSFAB-13] My supervisors [R]	Slightly approve = 40
[QSFAB-14] Other shift workers [R]	Approve = 20
	Strongly approve = 0
NB-scale – NB two	
What would the following people think if you were very fatigued mentally or physically while at work?	[R] = reverse scoring
[QSFAB-15] People your age [R]	Strongly disapprove = 100
[QSFAB-16] Co-workers [R]	Disapprove = 80
[QSFAB-17] Other healthcare workers [R]	Slightly disapprove = 60
[QSFAB-18] My supervisors [R]	Slightly approve = 40
[QSFAB-19] Other shift workers [R]	Approve = 20
	Strongly approve = 0
NB-scale scoring: summary scores for NB one and NB two is the sum of item scores divided by total number of items. The composite NB scale score is the sum of NB one and NB two divided by two	
NB-scale score interpretation: higher scores indicate a person believes the social norms and beliefs of his/her social network possess a negative view of behaviors that places an individual at work while very sleepy or fatigued	
SE-scale	0 = 0 – cannot do at all
Rate degree of confidence from 0 to 100 = cannot do at all–to–10 = highly certain can do	1 = 10
[QSFAB-20] Drink a coffee when sleepy or fatigued (__)	2 = 20
[QSFAB-21] Walk, jog, or run on duty to fight sleepiness or fatigue (__)	3 = 30
	4 = 40
	5 = 50 – moderately can do
	6 = 60
	7 = 70
	8 = 80
	9 = 90
	10 = 100 – highly certain can do
SE-scale scoring: summary score is the sum of item scores divided by total number of items	
SE-scale score interpretation: higher scores indicate the individual has a high-level of self-confidence he/she can perform select behaviors that may improve alertness and reduce feelings of sleepiness or fatigue while at work	
KNOW-scale – KNOW one	
Fatigue and sleepiness at work increase my risk of	
[QSFAB-22] Making a medical error	Strongly disagree = 0
[QSFAB-23] Being injured	Disagree = 20
[QSFAB-24] Being involved in an ambulance crash	Slightly disagree = 40
[QSFAB-25] Making a medication error	Slightly agree = 60
[QSFAB-26] Making a mistake	Agree = 80
[QSFAB-27] Losing the ability to concentrate	Strongly agree = 100
KNOW-scale – KNOW two	Strongly disagree = 0
[QSFAB-28] The hazards of fatigue and sleepiness on duty have been clearly demonstrated.	Disagree = 20
[QSFAB-29] Lack of sleep at home increase the risk of a fatigue-related error or injury while at work	Slightly disagree = 40
[QSFAB-30] The benefits of getting adequate sleep before shift work have been clearly demonstrated.	Slightly agree = 60
	Agree = 80
	Strongly agree = 100
KNOW-scale scoring: summary scores for KNOW one and KNOW two the sum of item scores divided by total items. The composite score equals the sum of KNOW one and KNOW two divided by two	
KNOW-scale score interpretation: higher scores indicate an individual has a high-level of awareness for the negative effects of sleepiness and fatigue while at work, that may be attributed to the acquisition of information, an increased understanding, or through experiences or education	

Table 5 (*Continued*)

Scale/sub-scale/items	Likert response scale and scoring
IMPORT-scale [QSFAB-31] The most important thing I do during my shift work is to maintain alertness and reduce fatigue. [QSFAB-32] I am willing to make sacrifices while at work to reduce feelings of fatigue and maintain alertness. [QSFAB-33] I would do whatever is necessary to reduce feelings of fatigue while at work and improve my alertness on duty.	Strongly disagree = 0 Disagree = 20 Slightly disagree = 40 Slightly agree = 60 Agree = 80 Strongly agree = 100
IMPORT-scale scoring: summary score for IMPORT scale is the sum of item scores divided by total items IMPORT-scale score interpretation: higher scores indicate an individual places a high-level of importance on the need to maintain alertness and reduce feelings of fatigue and/or sleepiness while at work	
EC-scale – EC one There are various reasons that might limit your ability to reduce feelings of fatigue and sleepiness while on duty. Please rate the importance of each listed when working at your primary EMS job. [QSFAB-34] Lack of interest from management [QSFAB-35] Lack of company policies that address this issue [QSFAB-36] My employer is not interested [QSFAB-37] Other issues are more important to the company [QSFAB-38] My co-workers and I are not given time to do things that reduce fatigue or sleepiness	Not at all important = 0 Somewhat important = 33.33 Important = 66.66 Very important = 100
EC-scale – EC two [QSFAB-39] Reducing fatigue or sleepy workers is a priority of my supervisors/company administrators [QSFAB-40] My company/employer has policies or procedures that address fatigue or sleepiness	Strongly disagree = 0 Disagree = 20 Slightly disagree = 40 Slightly agree = 60 Agree = 80 Strongly agree = 100
EC-scale – EC three [QSFAB-41] My responsibilities at home prevent me from doing things to reduce my fatigue or sleepiness when I'm at work. [QSFAB-42] My obligations to my family inhibit me from getting the sleep I need to recover between my shifts [QSFAB-43] I have multiple jobs that prevent me from getting the rest I need between shifts. [QSFAB-44] I have too many competing commitments and do not have time to get adequate sleep when at home and reduce feelings or sleepiness while at work	Strongly disagree = 0 Disagree = 20 Slightly disagree = 40 Slightly agree = 60 Agree = 80 Strongly agree = 100
EC-scale scoring: summary scores for EC one, EC two, and EC three scales is the sum of item scores divided by total items. The composite score equals the sum of EC one, EC two, and EC three divided by three EC-scale score interpretation: higher scores indicate the individual perceives his/her employer's policies and organizational related procedures/protocols (measured by EC one and EC two), or factors related to responsibilities unrelated to the organization (e.g., family duties as measured by EC three) as factors that inhibit the individual's ability to engage in behaviors that can improve alertness and reduce feelings of sleepiness or fatigue while at work	
HABIT scale [QSFAB-45] I frequently show up to work feeling fatigued [R] [QSFAB-46] Working while fatigued is something I do without thinking about it [R] [QSFAB-47] I work while feeling fatigued because I do this all the time [R]	[R] = reverse scoring Strongly disagree = 100 Disagree = 80 Slightly disagree = 60 Slightly agree = 40 Agree = 20 Strongly agree = 0
HABIT-scale scoring: summary score for HABIT scale is the sum of item scores divided by total items HABIT-scale score interpretation: higher scores indicate an individual has a habit of engaging in behaviors that may promote improved alertness and reduced feelings of sleepiness or fatigue while at work	
INTENT-scale [QSFAB-48] I am trying to improve my alertness at work and avoid feeling sleepy or fatigued [QSFAB-49] I do not plan to do anything to reduce my feeling sleepy on duty [R] [QSFAB-50] I do not plan to do anything while on duty to improve or maintain alertness [R]	[R] = reverse scoring for [QSFAB-& QSAFB#, QSFAB# positive scoring] Strongly disagree = 100 Disagree = 80 Slightly disagree = 40 Slightly agree = 60 Agree = 20 Strongly agree = 0
INTENT-scale scoring: summary score for INTENT scale is the sum of item scores divided by total items INTENT-scale score interpretation: higher scores indicate an individual's intent to engage in behaviors that may promote improved alertness and reduced feelings of sleepiness or fatigue while at work	

impact the validity of our findings. Emergency clinicians in other settings outside of our targeted area may have responded differently. One-fifth of our sample includes volunteer EMS workers. Our prior research shows that a greater proportion of full-time EMS workers report poor sleep and fatigue than part-time and volunteer (Patterson et al., 2012a). Responses to survey items may be affected by employment status. Despite this observation, we believe the percentage of part-time and volunteer EMS workers in this study is a strength. The vast majority of the >19,000 U.S. based EMS organizations are located in rural areas and primarily rely on a volunteer EMS workforce (Chapman et al., 2008; Mears et al., 2011). Greater numbers of volunteer EMS workers improve the representativeness of our study sample with the rest of the U.S. EMS workforce. We did not obtain measures of education on EMS worker participants. Some may be concerned that conference attendees are different in terms of education obtainment than non-attendees. It is possible that those attending regional conferences differ in terms of education and other factors. We believe this is a minor issue for our study given that the state of Pennsylvania has uniform criteria for EMS workers obtaining continuing education credits through multiple sources. EMS workers of all levels of certification must

meet a benchmark for continuing education, which includes attending regional conferences, didactic and clinical skills training, as well as online asynchronous education. Despite concerns over sampling, we believe any differences between our study sample and others are minor. A comparison of our study sample with other studies is supportive (Table 2). In terms of sample size, the size of study sample permitted tests of factor structure and reliability, and exceeds those in prior studies testing the psychometrics of pilot sleep/fatigue instruments (Buysse et al., 1989; Johns, 1991, 1992).

Our approach used classic psychometric techniques, such as assessment of content validity, factor structure analysis, and internal consistency, but we recognize that other approaches, such as item response theory analysis may also be useful. We adhered to one established guideline for scale development and began with identifying a well-established theoretical model (DeVellis, 1991). We adapted items from previously published scales and generated new items to measure multiple constructs of the IMBP. Numerous candidate items were generated for each scale tested. As recommended (DeVellis, 1991), we generated multiple draft items to create redundancy and a simple factor structure. We believe our approach stands apart from numerous published studies of sleep health and behavior instruments because many lack formal content validation (Knowlden et al., 2012; Robbins and Niederdeppe, 2014). We included validation of item content early to confirm or invalidate candidate items. It is important to remember that while expert opinions of content validity meet recommendations for instrument development and provide valuable information about item representativeness, they are subjective and subject to bias (Norbeck, 1985; Rubio et al., 2003).

We used established tests of validity and reliability to guide which items to retain or exclude. Tests of construct validity provided mixed findings. Tests of model fit from CFA analyses show that one or more indicators of model fit may be unacceptable for several scales. Tests of the IMPORT scale, EC sub-scales, HABIT scale, and INTENT scale revealed an unacceptable NNFI fit statistic. Bentler described the NNFI as a useful indicator when dealing with small sample sizes, yet cautioned use of “rules of thumb” or cut-points for this specific indicator because it may behave differently than other indices of model fit (Bentler, 1990). Numerous indicators exist for assessing model fit for the CFA procedure and there appears to be uncertainty in the literature regarding exactly which indicator(s) to report. For example, the chi-square statistic produced by the ‘proc calis’ procedure in SAS (Cary, North Carolina) is a commonly reported indicator of model fit. We did not report this indicator because, while it is widely used, the chi-square statistic is affected by sample size and other factors (Hatcher, 1994). Regardless of the indices reported, experts suggest any single indicator be accompanied with other indices and values be interpreted with caution (Bentler, 1990; Hatcher, 1994). It is ideal to have all indicators of model fit return acceptable values, yet it is apparent that analyses of model fit do not require that each indicator exceed proposed benchmarks (Hatcher, 1994). A mix of indicators is important for determining model fit. Investigators should be concerned about construct validity when none of the mix chosen return acceptable values (Hatcher, 1994). We reported multiple indicators of model fit to address uncertainty and limitations surrounding CFA testing. In this study, most indicators exceeded proposed benchmarks, signifying acceptable model fit.

Factor analyses of the self-efficacy scale did not produce a simple factor structure for CFA testing and traditional tests of internal consistency. Self-efficacy is concerned with an individual's perceived capability to perform a given task and items are traditionally phrased with “can do,” “will do,” “will,” or “can.” Self-efficacy is considered a direct and strong influential factor in the behavioral prediction. Bandura reports that there is no one-size fit all measure of self-efficacy and that each scale (set of items) must be calibrated for specific behaviors/tasks (Bandura, 2006). Nine of 13 candidate items did not meet the benchmark for content validity, leaving only 4 items for factor tests. Lack of items with redundancy contributed to our failure to identify a simple factor structure. We will follow up this study with a new effort for development and testing of a self-efficacy scale with a larger sample of EMS workers and sleep scientists. In the interim, we consider the two remaining items reported in Table 4 as preliminary measures of self-efficacy in the SFAB.

6. Conclusions

Measurement of sleep, fatigue, and alertness behavior among emergency medicine clinicians is challenging. We describe the development and psychometric testing of a pilot sleep, fatigue, and alertness behavior survey tool. Initial tests of validity, scale factors, and internal consistency support its use in future studies of EMS workers.

Author contribution statement

Drs. Daniel Patterson (PDP), Brian Suffoletto (BS), Daniel Buysse (DJB), and Mr. Weaver (MDW) conceived the study idea and framework. Dr. Patterson (PDP) and Mr. Matthew Weaver (MDW) developed the study protocol, data collection tool, study methodology, and carried out the collection of data. Dr. Patterson (PDP), Kyle McManigle (KLM), and Mr. Matthew Weaver (MDW) analyzed study data. All authors/investigators examined study findings, interpreted study findings, and synthesized study findings for reporting in a peer-reviewed manuscript. All authors participated in development and editing of the draft manuscript, providing input and editing for all components, including the introduction, methods, results, discussion and conclusions.

Competing interests

The authors report no competing interests or conflicts of interests.

Acknowledgements

Work on this study was supported by Dr. Patterson's KL2 training grant from the National Center for Research Resources and the National Institutes of Health (NIH/NCATS grant no: KL2 TR000146 (Dr. Reis PI). The conclusions, views, opinions, and content in this paper should not be interpreted as reflecting the opinions of the NIH.

References

Akerstedt, T., Gillberg, M., 1990. Subjective and objecive sleepiness in the active individual. Int. J. Neurosci. 52 (1–2), 29–37.
Allen, S., Ortlepp, K., 2002. Conceptualising and operationalising work versus career salience. SA J. Ind. Psychol. 28 (2), 7–14.
Anderson, J.C., Gerbing, D.W., 1988. Structual equation modeling in practice: a review and recommended two-step approach. Psychol. Bull. 103 (3), 411–423.
Bandura, A., 2006. Guide for constructing self-efficacy scales. In: Pajares, F., Urdan, T. (Eds.), Self-efficacy Beliefs of Adolescents. Information Age Publishing, Greenwich, CT, pp. 307–337.
Bauder, B., 2012. Safety Fears Rise With City's Use of Overtime. Pittsburgh Tribune-Review, Pittsburgh.
Belenky, G., Akerstedt, T., 2011. Occupational sleep medicine, In: Kryger, M.H., Roth, T., Dement, W.C. (Eds.), Principles and Practice of Sleep Medicine. 5th ed. Elsevier, St. Louis, MI, pp. 734–737.
Bentler, P.M., 1990. Comparative fit indexes in structual models. Psychol. Bull. 107 (2), 238–246.
Borbely, A.A., 1982. A two process model of sleep regulation. Hum. Neurobiol. 1 (3), 195–204.

Boudreaux, E., Mandry, C., Brantley, P.J., 1998. Emergency medical technician schedule modification: impact and implications during short and long-term follow-up. Acad. Emerg. Med. 5 (2), 128–133.
Bowman, B., Broyles, G., Bryner, N., Burgess, J., Caplan, M., Duffy, R., Gelberg, K., Grant, C., Johnson, W., Maguire, B., Montgomery, B., Neimiller, R., Nenneman, M., Schultz, E., Seabury, S., Sommers, K., Stagnaro, V., Thompson, J., Vermilye, E., Wijetunge, G., Wilson, M., D'Allessandro, M., Haskell, W., Butler, C., Palcic, J., Charles, L., Daniels, R., Frederick, L., Funk, R., Grubb, P., Hales, T., Hornsby-Meyers, J., Kubale, T., Moore, P., Reissman, D., 2013. National Occupational Research Agenda For Occupational Safety and Health Research and Practice in the U. S. Public Safety Sector, vol. 2013. Center for Disease Control and Prevention, NIOSH, United States.
Brice, J.H., Studnek, J.R., Bighma, B.L., Martin-Gill, C., Custalow, C.B., Hawkins, E., Morrison, L.J., 2012. EMS provider and patient safety during response and transport: proceedings of an ambulance safety conference. Prehosp. Emerg. Care 16 (1), 3–19.
Brown, W., Dickison, P., Misselbeck, W., Levine, R., 2002. Longitudinal emergency medical technician attribute and demographic study (LEADS): an interim report. Prehosp. Emerg. Care 6 (4), 433–439.
Bureau Of Labor Statistics, 2014. USD (2014, Wednesday, January 8, 2014). Occupational Outlook Handbook, 2014–2015 Edition, EMTs and Paramedics., from http://www.bls.gov/ooh/healthcare/emts-and-paramedics.htm (retrieved 09.01.14).
Buysse, D.J., Reynolds, C.F.R., Monk, T.H., Berman, S.R., Kupfer, D.J., 1989. The Pittsburgh sleep quality index: a new instrument for psychiatric practice and research. Psychiatry Res. 28 (2), 193–213.
Caruso, C.C., Bushnell, T., Eggerth, D., Heitmann, A., Kojola, B., Newman, K., Rosa, R.R., Sauter, S.L., Vila, B., 2006. Long working hours, safety, and health: toward a National Research Agenda. Am. J. Ind. Med. 49 (11), 930–942.
Caruso, C.C., Hitchcock, E.M., Dick, R.B., Russo, J.M., Schmit, J.M., 2004. Overtime and Extended Work Shifts: Recent Findings on Illnesses, Injuries, and Health Behaviors. National Institute for Occupational Safety and Health, Cincinnati, OH.
Chalder, T., Berelowitz, G., Pawlikowska, T., Watts, L., Wessely, S., Wright, D., Wallace, E.P., 1993. Development of a fatigue scale. J. Psychosom. Res. 37 (2), 147–153.
Chapman, S.A., Lindler, V., Kaiser, J.A., Nielsen, C.S., Bates, T., Hailer-O'Keefe, L., 2008. Emergency Medical Services Workforce for the 21st Century Project. National Highway Traffic Safety Administration, Washington D.C.
Daan, S., Beersma, D.G., Borbely, A.A., 1984. Timing of human sleep: recovery process gated by a circadian pacemaker. Am. J. Physiol. 246 (2 Pt. 2), R161–R183.
Davis, D.P., Garberson, L.A., Andrusiek, D.L., Hostler, D., Daya, M., Pirallo, R., Craig, A., Stephens, S., Larsen, J., Drum, A.F., Fowler, R., 2007. A descriptive analysis of Emergency Medical Service Systems participating in the Resuscitation Outcomes Consortium (ROC) network. Prehosp. Emerg. Care 11 (4), 369–382.
Dawson, D., Chapman, J., Thomas, M.J., 2012. Fatigue-proofing: a new approach to reducing fatigue-related risk using the principles of error management. Sleep Med. Rev. 16 (2), 167–175.
De Pelsmacker, P., Janssens, W., 2007. The effect of norms, attitudes and habits on speeding behavior: scale development and model building and estimation. Accid. Anal. Prev. 39 (1), 6–15.
Delucchi, K.L., Tajima, B., Guydish, J., 2009. Development of the smoking knowledge, attitudes, and practices (S-KAP) instrument. J. Drug Issues 39 (2), 347–364.
DeVellis, R.F., 1991. Scale Development: Theory and Applications, 26. SAGE Publications, Newbury Park, CA.
Drucker, C., Gerberich, S.G., Manser, M.P., Alexander, B.H., Chruch, T.R., Ryan, A.D., Becic, E., 2013. Factors associated with civilian drivers involved in crashes with emergency vehicles. Accid. Anal. Prev. 55, 116–123.
Dunham, R.B., Pierce, J.L., 1986. Attitudes toward work schedules: construct definition, instrument development, and validation. Acad. Mange. J. 29 (1), 170–182.
Eisinga, R., Grotenhuis, M., Pelzer, B., 2013. The reliability of a two-item scale: Pearson Cronbach, or Spearman–Brown? Int. J. Public Health 58 (4), 637–642.
Fishbein, M., Yzer, M.C., 2003. Using theory to design effective health behavior interventions. Commun. Theory 13 (2), 164–183.
Frakes, M.A., Kelly, J.G., 2007. Sleep debt and outside employment patterns in helicopter air medical staff working 24-hour shifts. Air Med. J. 26 (1), 45–49.
Ganley, B.J., Rosario, D.I., 2013. The smoking attitudes, knowledge, intent, and behaviors of adolescents and young adults: implications for nursing practice. J. Nurs. Educ. Pract. 3 (1), 40–50. doi:http://dx.doi.org/10.5430/jnep.v3n1p40.
Greene, M., 2013. Living the life in EMS. JEMS 38 (10), 34–41.
Griffeth, R.W., Hom, P.W., Gaertner, S., 2000. A meta-analysis of antecedents and correlates of employee turnover: update, moderator tests, and research implications for the next millennium. J. Manage. 26 (3), 463–488.
Hatcher, L., 1994. A Step-by-Step Approach to Using the SAS System for Factor Analysis and Structural Equation Modeling. SAS Institute, Inc., Cary, NC.
International Agency Research on Cancer (IARC), 2010. Shiftwork. In: World Health Organization (Ed.), The IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. International Agency for Research on Cancer, Lyon, France and Geneva, Switzerland.
Johns, M.W., 1991. A new method for measuring daytime sleepiness: the Epworth Sleepiness Scale. Sleep 14 (6), 540–545.
Johns, M.W., 1992. Reliability and factor analysis of the Epworth Sleepiness Scale. Sleep 15 (4), 376–381.
Kirkwood, S., 2012. Studies prove 24-hour shifts unsafe: another ethical dilemma for EMS leaders. JEMS (April) .
Knowlden, A.P., Sharma, M., Bernard, A.L., 2012. A theory of planned behavior research model for predicting the sleep intentions and behaviors of undergraduate college students. J. Prim. Prev. 33 (1), 19–31.
Lance, C.E., Butts, M.M., Michels, L.C., 2006. The sources of four commonly reported cutoff criteria: what did they really say? Organ. Res. Methods 9 (2), 202–220.
Lynn, M.R., 1986. Determination and quantification of content validity. Nurs. Res. 35 (6), 382–385.
Marsh, H.W., Balla, J.R., McDonald, R.P., 1988. Goodness-of-fit indexes in confirmatory factor-analysis: the effect of sample-size. Psychol. Bull. 103 (3), 391–410.
Mears, G., Armstrong, B., Fernandez, A.R., Mann, N.C., McGinnis, K., Mears, C.R., Sanddal, N.D., Sanddal, T.L., Shofer, F.S., 2011. National EMS Assessment. Available from: http://www.ems.gov/pdf/2011/National_EMS_Assessment_Final_Draft_12202011.pdf.
National EMS Advisory Council, 2013. Fatigue in Emergency Medical Services. National EMS Advisory Council Quarterly Meeting. From: http://ems.gov/pdf/nemsac/2013/NEMSAC-AdvisoryFatigueJan2013.pdf (30.01.2013).
Nevo, B., 1985. Face validity revisited. J. Educ. Meas. 22 (4), 287–293.
National Highway Traffic Safety Administration's Office of EMS, 2013. Strategy for a National EMS Culture of Safety. National Highway Traffic Safety Administration Office of EMS, United States.
Norbeck, J.S., 1985. What constitutes a publishable report of instrument development? Nurs. Res. 34 (6), 380–382.
Olds, R.S., Thombs, D.L., Tomasek, J.R., 2005. Relations between normative beliefs and initiation intentions toward cigarette, alcohol, and marijuana. J. Adolesc. Health 37 (1), 75 e77–75. e13.
Patterson, P.D., Huang, D.T., Fairbanks, R.J., Simeone, S.J., Weaver, M.D., Wang, H.E., 2010a. Variation in emergency medical services workplace safety culture. Prehosp. Emerg. Care 14 (4), 448–460.
Patterson, P.D., Huang, D.T., Fairbanks, R.J., Wang, H.E., 2010b. The emergency medical services safety attitudes questionnaire. Am. J. Med. Qual. 25 (2), 109–115.
Patterson, P.D., Suffoletto, B.P., Kupas, D.F., Weaver, M.D., Hostler, D., 2010c. Sleep quality and fatigue among prehospital providers. Prehosp. Emerg. Care 14 (2), 187–193.
Patterson, P.D., Weaver, M.D., Frank, R.C., Warner, C.W., Martin-Gill, C., Guyette, F.X., Fairbanks, R.J., Hubble, M.W., Songer, T.J., Callaway, C.W., Kelsey, S.F., Hostler, D., 2012a. Association between poor sleep, fatigue, and safety outcomes in emergency medical services providers. Prehosp. Emerg. Care 16 (1), 86–97.
Patterson, P.D., Weaver, M.D., Hostler, D., Guyette, F.X., Callaway, C.W., Yealy, D.M., 2012b. The shift length, fatigue, and safety conundrum in EMS. Prehosp. Emerg. Care 16 (4), 572–576.
Patterson, P.D., Weaver, M.D., Weaver, S.J., Rosen, M.A., Todorova, G., Weingart, L.R., Krackhardt, D., Lave, J.R., Arnold, R.M., Yealy, D.M., Salas, E., 2012c. Measuring teamwork and conflict among emergency medical technician personnel. Prehosp. Emerg. Care 16 (1), 98–108.
Pirrallo, R.G., Loomis, C.C., Levine, R., Woodson, B.T., 2012. The prevalence of sleep problems in emergency medical technicians. Sleep Breath 16 (1), 149–162.
Polit, D.F., Beck, C.T., Owen, S.V., 2007. Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. Res. Nurs. Health 30 (4), 459–467.
Ray, A.M., Kupas, D.F., 2007. Comparison of rural and urban ambulance crashes in Pennsylvania. Prehosp. Emerg. Care 11 (4), 416–420.
Ream, E., Richardson, A., 1996. Fatigue: a concept analysis. Int. J. Nurs. Stud. 33 (5), 519–529.
Reichard, A.A., Marsh, S.M., Moore, P.H., 2011. Fatal and nonfatal injuries among emergency medical technicians and paramedics. Prehosp. Emerg. Care 15 (4), 511–517.
Rise, J., Kovac, V., Kraft, P., Moan, I.S., 2008. Predicting the intention to quit smoking and quitting behavior: extending the theory of planned behavior. Br. J. Health Psychol. 13 (Pt 2), 291–310.
Robbins, R., Niederdeppe, J., 2014. Using the integrative model of behavioral prediction to identify promising message strategies to promote healthy sleep behavior among college students. Health Commun.
Rubio, D.M., 2005. Content validity. In: Kempf-Leonard, K. (Ed.), Encyclopedia of Social Measurement. Academic Press, Burlington, MA, pp. 495–498.
Rubio, D.M., Berg-Weger, M., Tebb, S.S., Lee, E.S., Rauch, S., 2003. Objectifying content validity: conducting a content validity study in social work research. Soc. Work Res. 27 (2), 94–104.
Takeyama, H., Itani, T., Tachi, N., Sakamura, O., Murata, K., Inoue, T., Takanishi, T., Suzumura, H., Niwa, S., 2009. Effects of a modified ambulance night shift system on fatigue and physiological function among ambulance paramedics. J. Occup. Health 51 (3), 204–209.
Ware, J.E.J., Gandek, B., 1998. Methods for testing data quality, scaling assumptions, and reliability: the IQOLA project approach (international quality of life assessment). J. Clin. Epidemiol. 51 (11), 945–952.
Williamson, A., Lombardi, D.A., Folkard, S., Stutts, J., Courtney, T.K., Connor, J.L., 2011. The link between fatigue and safety. Accid. Anal. Prev. 43 (2), 498–515.
Winwood, P.C., Winefield, A.H., Dawson, D., Lushington, K., 2005. Development and validation of a scale to measure work-related fatigue and recovery: the Occupational Fatigue Exhaustion/Recovery Scale (OFER). J. Occup. Environ. Med. 47 (6), 594–606.